Deployment of Privacy-Preserving Machine Learning for Political Polling in the 2024 Presidential Election

Sam Buxbaum

Lucas M. Tassis, Lucas Boschelli, Giovanni Comarela, Mayank Varia, Mark Crovella, Dino P. Christenson

PPML Workshop

August 17, 2025

Sam Buxbaum

Lucas M. Tassis

Lucas Boschelli

Giovanni Comarela

Mayank Varia

Mark Crovella

Dino P. Christenson

Overview

- We build a system for securely predicting political preferences from web browsing data
- We collect and analyze data from almost 8000 unique users
- All analysis takes place under MPC

Roadmap

1. Motivation

2. System Design

3. Learning Algorithm

4. Lessons Learned and Future Directions

Motivation

Background

- Web browsing behavior can predict voting results
- Quantifying the 'Comey letter' (Comarela et al.)
  - The event was too close to the election for other polling methods to detect the effect

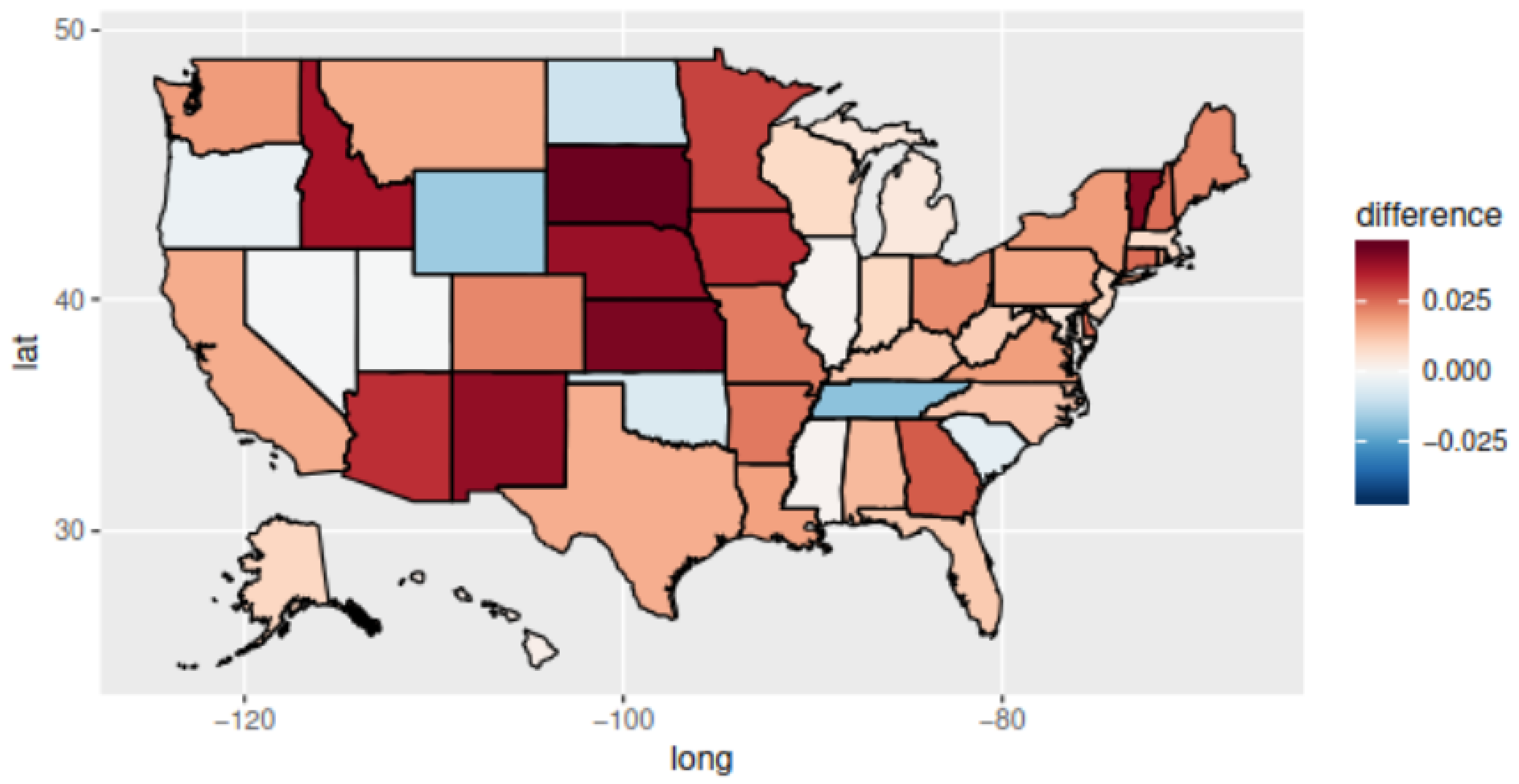


Figure 8: Impact of the 'Comey letter' at the state level.

Traditional Political Polling

- Data collection takes time
  - High latency between poll commission and results
- Human-intensive data collection
  - Scaling to collect much more data would be costly
- Poor geographic and temporal coverage
  - Results are concentrated in key regions immediately before an election
  - Many locations go unpolled, particularly early in an election cycle

West Virginia 2024 Presidential Election Polls

Harris vs. Trump

Source	Date	Sample	Harris	Trump	Other
Research America	8/30/2024	400 LV ±4.9%	34%	61%	5%

Michigan 2024 Presidential Election Polls

Instantly compare a poll to prior one by same pollster

Harris vs. Trump

Source	Date	Sample	Harris	Trump	Other
Average of 23 Polls†			48.6%	46.8%	-
FAU / Mainstreet	11/04/2024	713 LV	49%	47%	4%
Emerson College	11/04/2024	790 LV ±3.4%	50%	48%	2%
Research Co.	11/04/2024	450 LV ±4.6%	49%	47%	4%
InsiderAdvantage	11/03/2024	800 LV ±3.7%	47%	47%	6%
Trafalgar Group	11/03/2024	1,079 LV ±2.9%	47%	48%	5%
MIRS / Mich. News Source	11/03/2024	585 LV ±4%	50%	48%	2%
NY Times / Siena College	11/03/2024	998 LV ±3.7%	47%	47%	6%
Morning Consult	11/03/2024	1,108 LV ±3%	49%	48%	3%
AtlasIntel	11/02/2024	1,198 LV ±3%	48%	50%	2%
Redfield & Wilton	11/01/2024	1,731 LV ±2.2%	47%	47%	6%
The Times (UK) / YouGov	11/01/2024	942 LV ±3.9%	48%	45%	7%
EPIC-MRA	11/01/2024	600 LV ±4%	48%	45%	7%
Marist Poll	11/01/2024	1,214 LV ±3.5%	51%	48%	1%
AtlasIntel	10/31/2024	1,136 LV ±3%	49%	49%	2%
Echelon Insights	10/31/2024	600 LV ±4.4%	48%	48%	4%
MIRS / Mich. News Source	10/31/2024	1,117 LV ±2.5%	47%	49%	4%
UMass Lowell	10/31/2024	600 LV ±4.5%	49%	45%	6%
Washington Post	10/31/2024	1,003 LV ±3.7%	47%	46%	7%
Fox News	10/30/2024	988 LV ±3%	49%	49%	2%
CNN	10/30/2024	726 LV ±4.7%	48%	43%	9%
Suffolk University	10/30/2024	500 LV ±4.4%	47%	47%	6%

Two Approaches to Political Polling

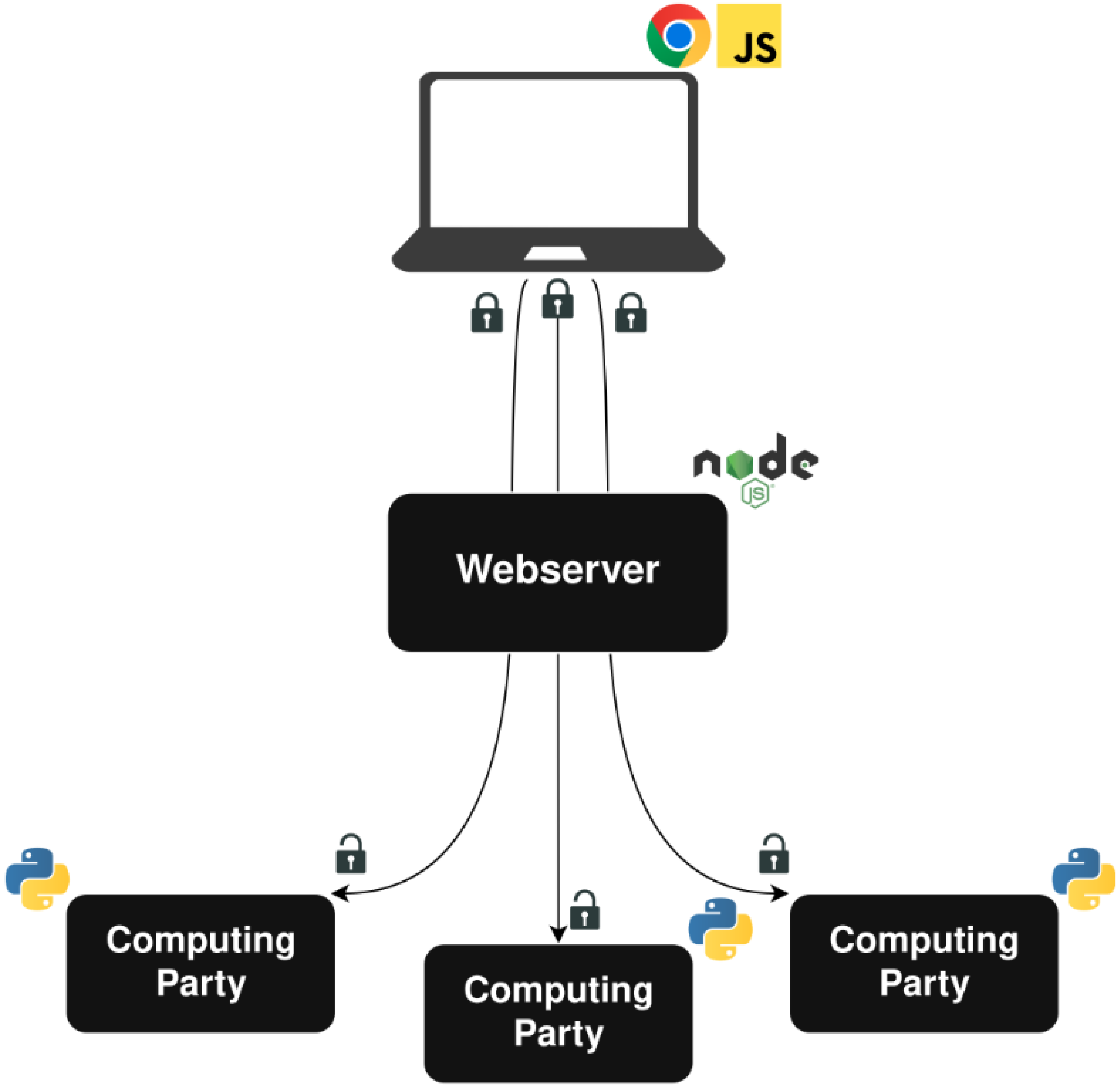
Traditional Polling		Web Behavior Analysis
Slow	VS	Immediate
Expensive		Cheap
Coarse-grained insights		Fine-grained insights

What about privacy?

System Design

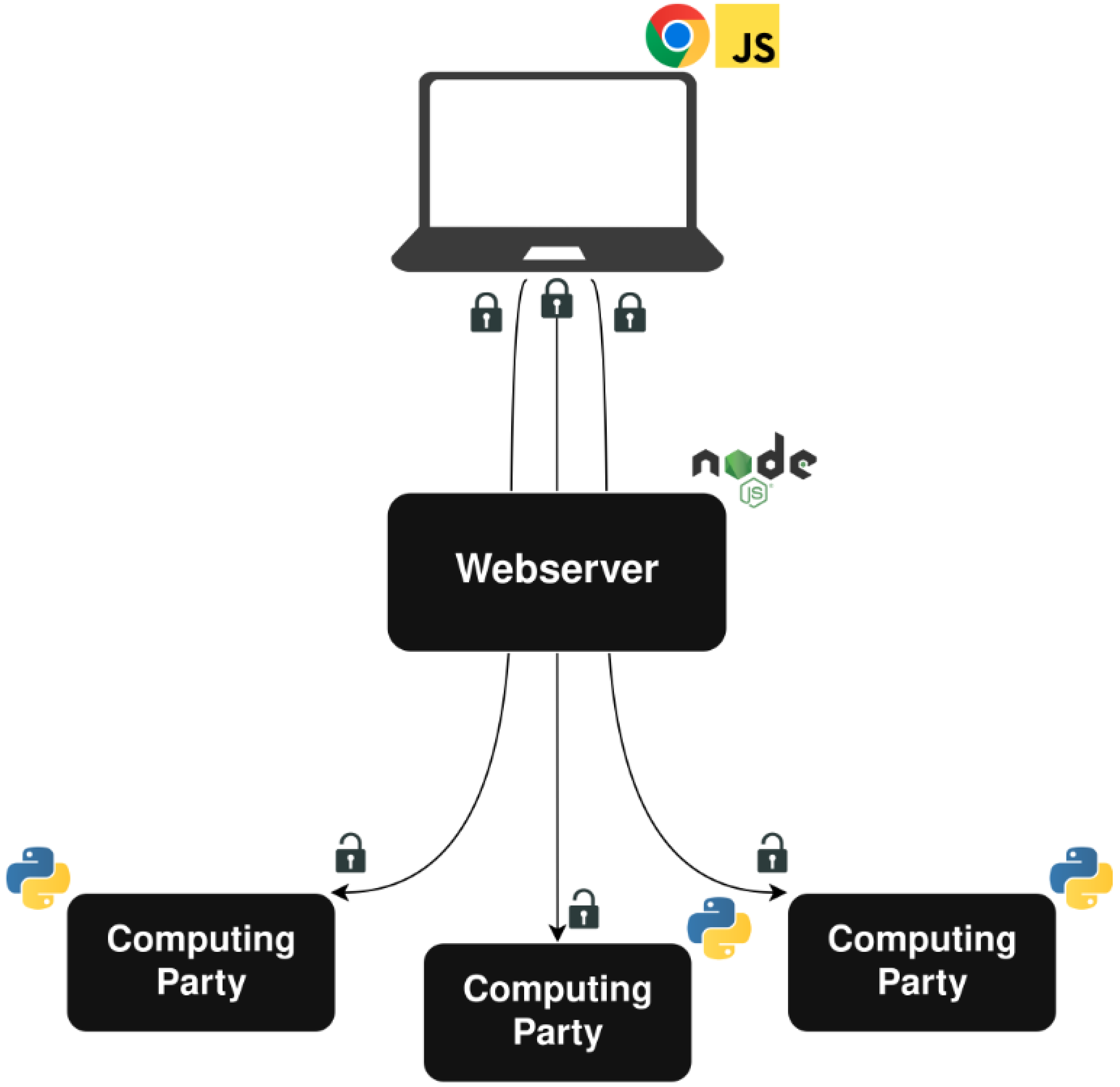
Three Components

- Client Plugin
- Webserver
- MPC Backend



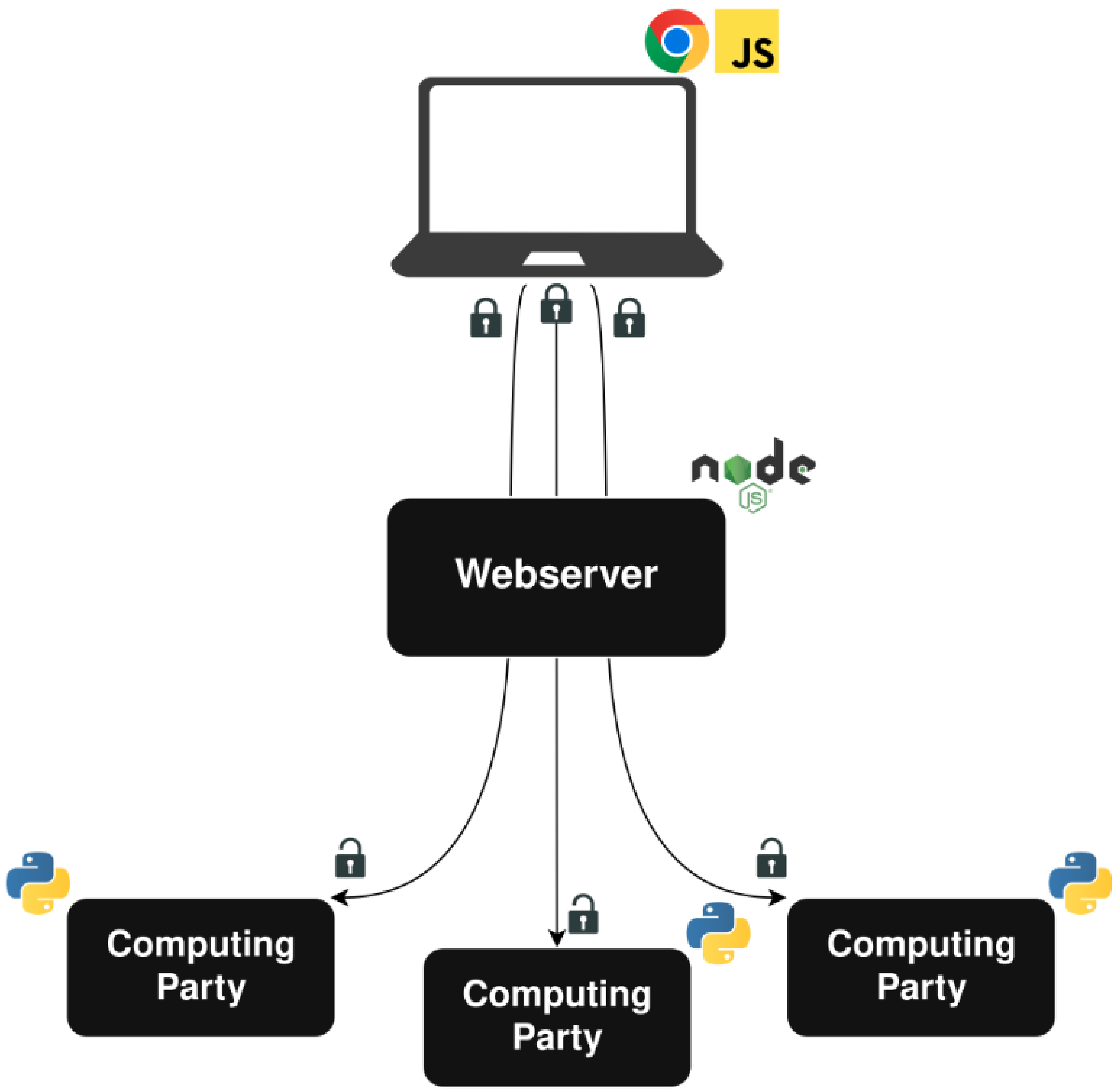
Client Plugin

- Custom-built Chrome plugin to monitor browsing
  - Tracks visits to top websites
  - Tracks referrals to top websites from social media sites
- Daily data uploads of secret-shared histograms
- Client-side secret sharing and encryption
- Implementation is open source



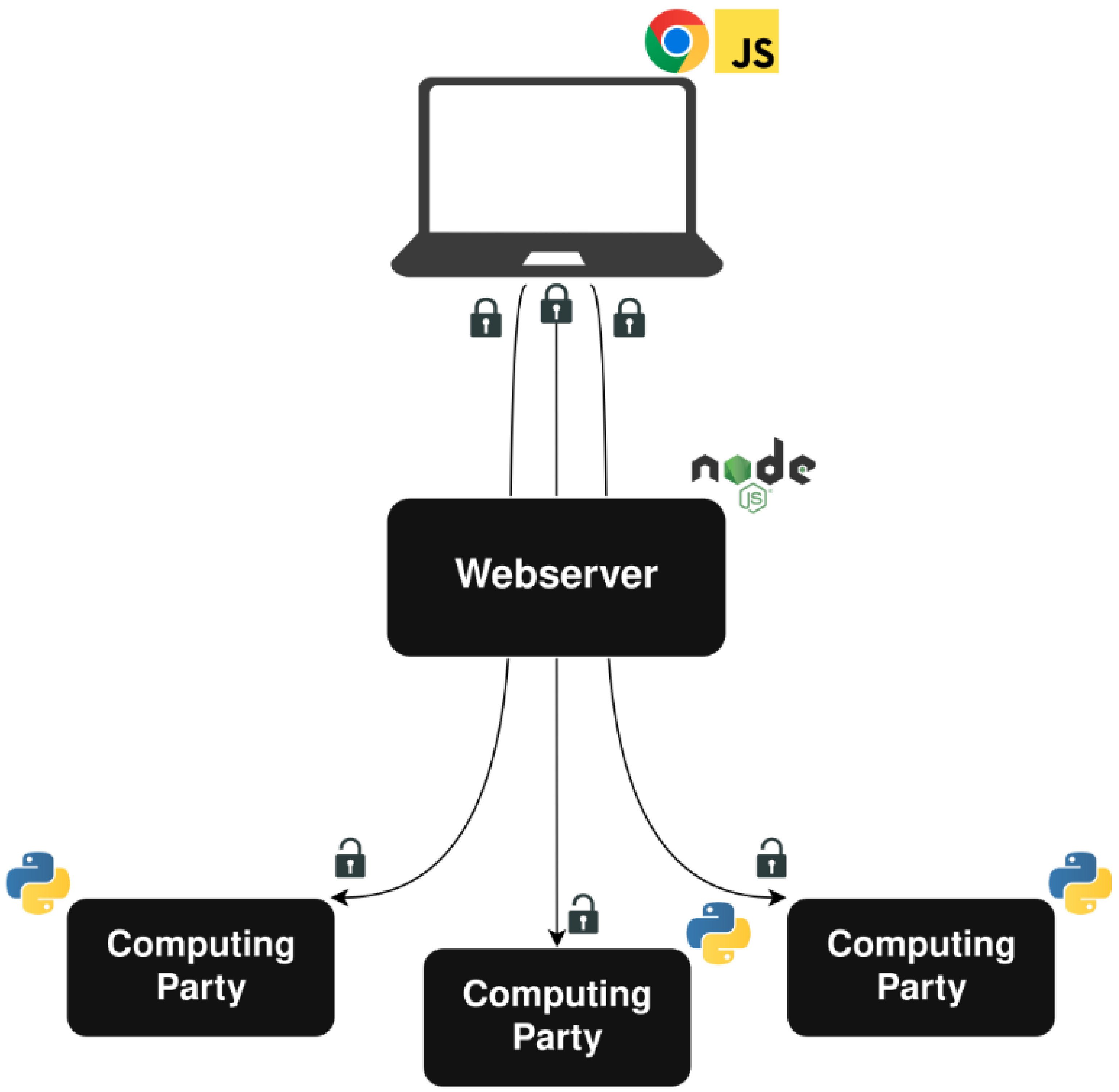
Webserver

- Simplifies interaction with clients
- Collects basic metadata
  - For payment and location tracking
- Never sees any private data
  - Secret-shares are end-to-end encrypted to the parties



MPC Backend

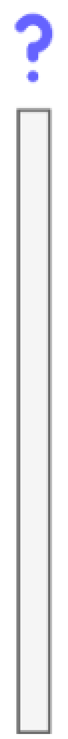
- Threat model
  - Three parties
  - Semi-honest security
  - One adversary
- We use and augment the CrypTen library
- Code will be available in the future

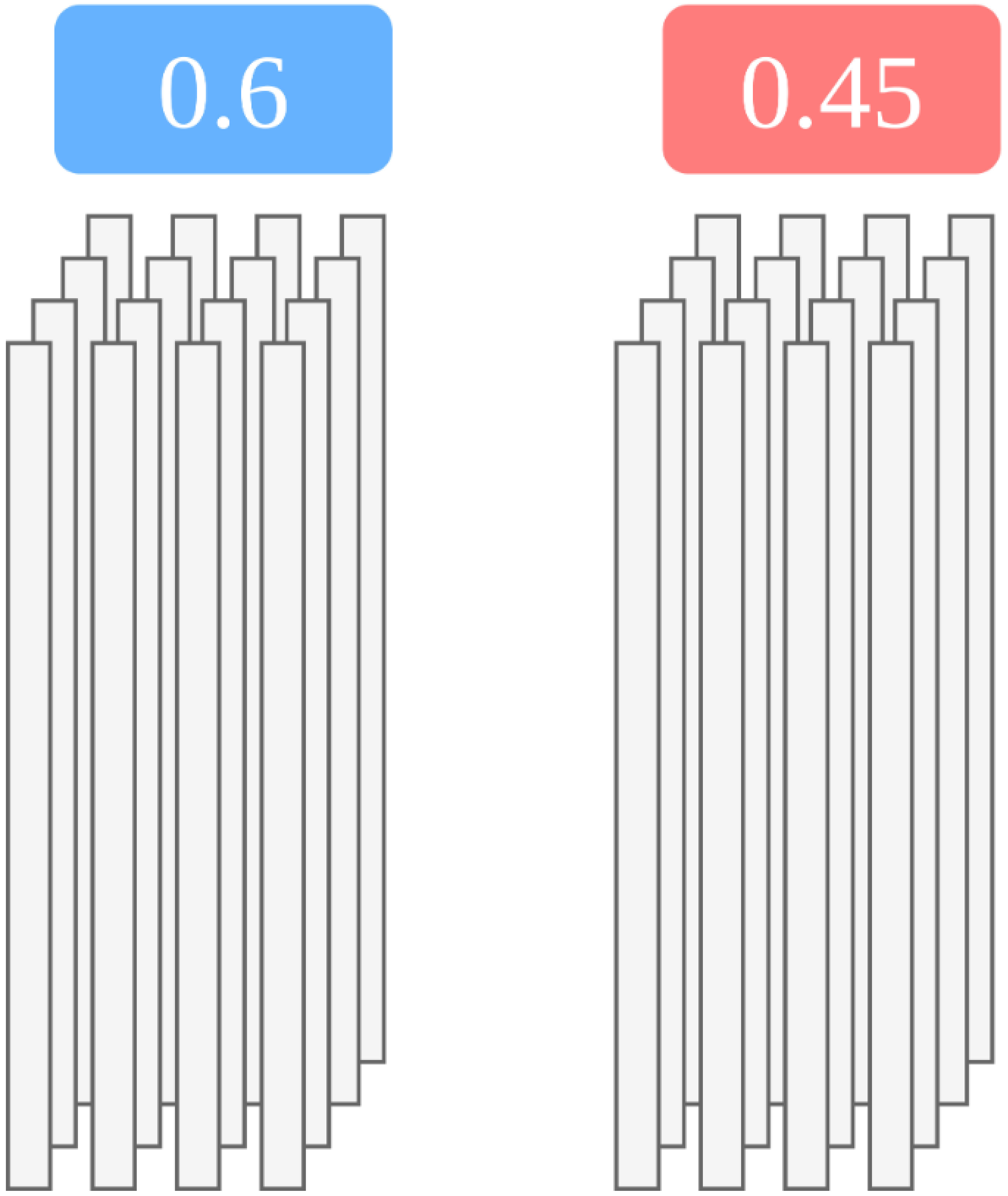


The Learning Process

Learning from Label Proportions (LLP)

- Each user uploads an *unlabeled* 1,034-element vector every day
  - Number of visits to the top 517 sites
  - Number of social media referrals to the top 517 sites
- Unlabeled vectors are grouped by state
- Each state has a ground-truth label
- Train on aggregate ground truth
- Predict on an individual level





The Learning Algorithm

Input: visit and referral histograms grouped by state, state-level ground truth

Train–update loop until convergence

- Assign initial predictions to each individual
- Train a logistic regression model on current predictions
- Compute new predictions
  - Sort each state's users by prediction
  - Set a threshold so the aggregate state prediction matches state ground truth
  - Assign updated individual predictions according to the threshold
- Repeat until predictions converge

Implementation in MPC

- Initial label assignment can be performed in plaintext
- Training a logistic regression model is implemented in CrypTen
  - Logistic regression is supported out-of-the-box
- Computing thresholds requires oblivious sorting
  - We implement Bitonic sort using CrypTen's secure primitives
  - The most expensive piece of the computation
- Updated label assignment and convergence checking use secure comparisons
- Practically efficient
  - All computation completed in a few hours at most

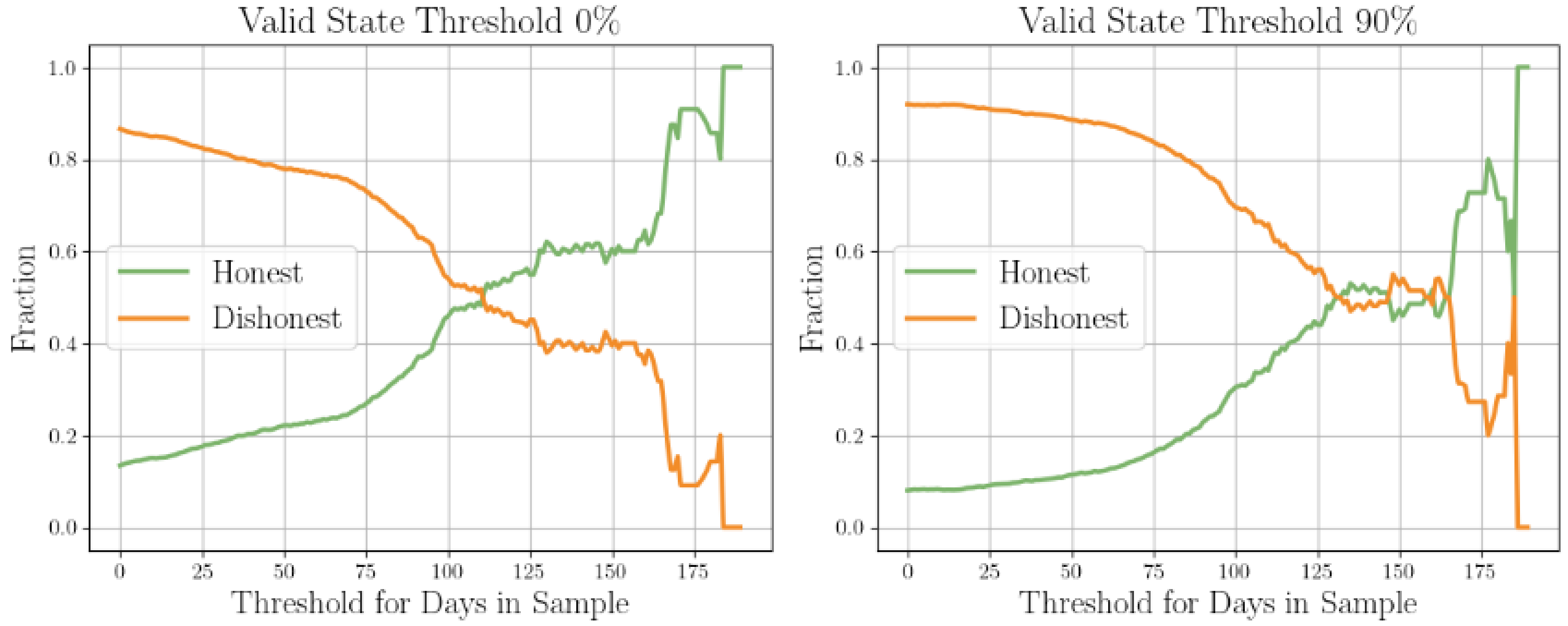
Lessons Learned and Future Directions

Data Integrity Matters

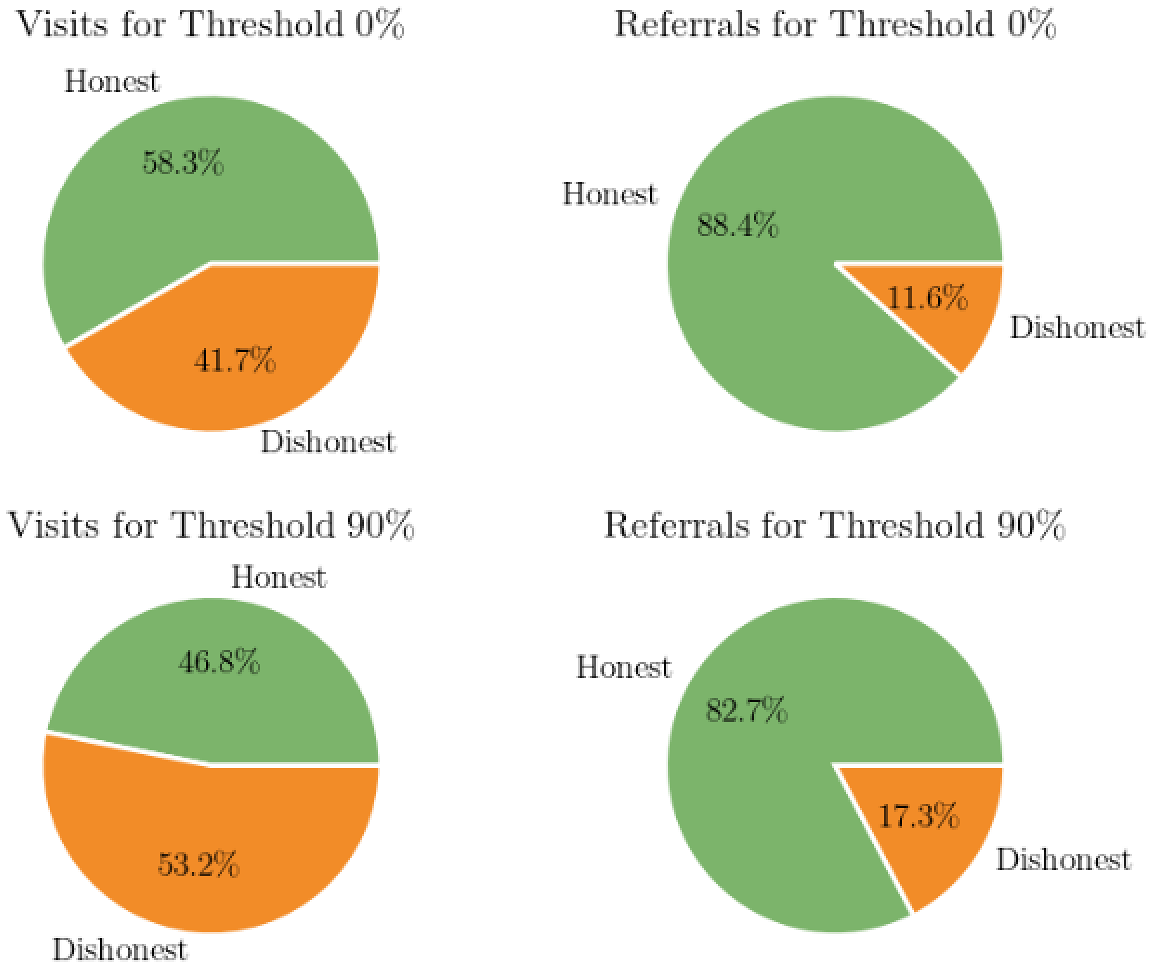
- Our initial advertisement targeted users in the swing states
  - Incentive to report location dishonestly
- We use IP addresses and geolocation data to validate the data
  - 98% of users are located in the United States
- State-level results are concerning
  - 85% of users reported their state dishonestly
  - 15% of users reported an invalid ZIP code for their state

Digging Deeper on the Data

- Users in the sample for longer are more honest
  - Ephemeral dishonest users, consistent honest users



- Honest users contribute much richer data
  - Referral data provides the best signal (Comarela et al.)



Data Integrity

Lesson: Validating and enforcing user honesty should be a priority in future deployments.

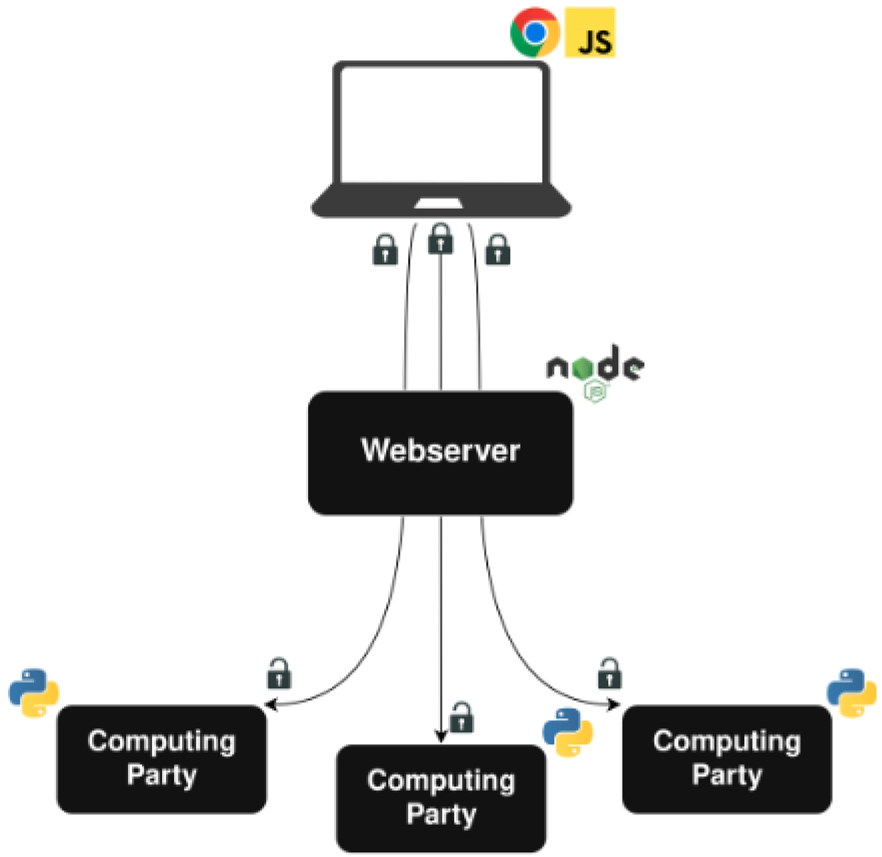
Lesson: Our learning process is surprisingly robust to dishonest users.

Opportunity — Privacy-Preserving Location Verification

- A difficult problem, even without privacy
- Integrate cryptographic techniques with existing plaintext approaches

Strengthening the Threat Model

- AWS as a single point of failure
- Reduce or eliminate trust in the core computation
  - Incorporate external organizations in the MPC
  - Explore different cryptographic primitives
- Anonymous payments



Thank You!

sambux@bu.edu